Journal of Computational Physics 228 (2009) 5298–5315

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

An exactly conservative particle method for one dimensional scalar conservation laws

Yossi Farjoun, Benjamin Seibold *

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA 02139, USA

ARTICLE INFO

Article history:
Received 4 September 2008
Received in revised form 7 April 2009
Accepted 9 April 2009
Available online 22 April 2009

MSC:
65M25
35L65

Keywords:
Conservation law
Meshfree
Particle method
Particle management

ABSTRACT

A particle scheme for scalar conservation laws in one space dimension is presented. Particles representing the solution are moved according to their characteristic velocities. Particle interaction is resolved locally, satisfying exact conservation of area. Shocks stay sharp and propagate at correct speeds, while rarefaction waves are created where appropriate. The method is variation diminishing, entropy decreasing, exactly conservative, and has no numerical dissipation away from shocks. Solutions, including the location of shocks, are approximated with second order accuracy. Source terms can be included. The method is compared to CLAWPACK in various examples, and found to yield a comparable or better accuracy for similar resolutions.

1. Introduction

Conservation laws are important models for the evolution of continuum quantities, describing shocks and rarefaction behavior. Fundamental mathematical properties are global and local conservation, the presence of similarity solutions, and the method of characteristics. Successful numerical methods employ these properties to their advantage: finite difference methods yield correct shock speeds if applied in conservation form. Finite volume methods are fundamentally based on conservation properties. Godunov schemes [10], front tracking methods [12], and many related approaches, approximate the global solution by local similarity solutions. The method of characteristics is used in the CIR method [4] in combination with an interpolation scheme. Although for scalar equations it provides a direct formula for the solution (where it is smooth), it is less popular, since it does not possess conservation properties. Consequently, basic CIR schemes do not yield correct shock speeds.

Many commonly used numerical methods operate on a fixed Eulerian grid. Advantages are simple data structures and an easy generalization to higher space dimensions. Eulerian schemes can be constructed by tracking the "correct" approximate solution for a short time step, either by solving local Riemann problems (Godunov [10]) or by tracing characteristics (CIR), followed by an interpolation step, at which the solution is remapped onto the fixed grid. This "remeshing" step generally yields numerical dissipation and dispersion. Since the shortest interaction time between shocks or characteristics

* Corresponding author.
E-mail addresses: yfarjoun@math.mit.edu (Y. Farjoun), seibold@math.mit.edu (B. Seibold).
URL: http://www-math.mit.edu/~seibold/ (B. Seibold).

0021-9991/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2009.04.013

determines the global time step, remeshing is performed unnecessarily in many places. In practice, Eulerian methods require sophisticated schemes to obtain solutions with sharp features, but without creating oscillation. Finite volume methods are equipped with limiters [20], while finite difference methods use nonlinear approximations, such as ENO [11] or WENO [16].

An alternative approach is to abandon the Eulerian property, and thus avoid remeshing. Godunov methods become front tracking methods, at least in one space dimension. While in the former the interaction of shocks is avoided by remeshing, in the latter it is resolved after approximating the flux function by a piecewise linear function. By construction, front tracking is successful when representing shocks, but cumbersome when approximating smooth parts of the solution. Similarly, CIR methods become Lagrangian particle methods. Particles carry function values and move with their characteristic velocities. As motivated in [8], this provides a simple and accurate solution method for conservation laws, without ever approximating derivatives. However, particle management is required, for two reasons: first, neighboring particles may depart from each other, resulting in poorly resolved regions. This is prevented by inserting particles into gaps. Second, particles may collide. If left unchecked, such a shock event leads to a "breaking wave" solution. This is prevented by merging particles upon collision.

Lagrangian particle methods have been successfully applied in the simulation of fluid flows. Examples are vortex methods [2], smoothed particle hydrodynamics (SPH) [17,9,18], or generalized SPH methods [6]. The solution is approximated on a cloud of points which move with the flow, thus the governing equations are discretized in their more natural Lagrangian frame of reference. In specific applications, more accurate solutions may be obtained than with fixed grid approaches. In addition, with particles local adaptivity is a straightforward extension.

The particle method presented here combines the method of characteristics (where the solution is smooth) and particle merges (at shocks). The evolution of area between neighboring particles is derived from local similarity solutions. The method is designed to conserve area exactly.

1.1. Formulation of the particle method

The simplest form of a one dimensional scalar conservation law is

$$u_t + f(u)_x = 0, \quad u(x,0) = u_0(x) \tag{1}$$

with f' continuous. The characteristic equations [7]

$$\begin{cases} \dot{x} = f'(u) \\ \dot{u} = 0 \end{cases} \tag{2}$$

yield the solution (while it is smooth) forward in time: at each point $(x_0, u_0(x_0))$ a characteristic curve $x(t) = x_0 + f'(u_0(x_0))t$ starts, carrying the function value $u(x(t),t) = u_0(x_0)$. While the particle method is presented here for the simple case (1), the method of characteristics applies in more general cases, such as space-dependent flux functions and source terms (see Section 7). When characteristic curves collide, a shock arises. It moves at a speed so that area (under the function $u(\cdot,t)$) evolves correctly with respect to (1). The Rankine–Hugoniot condition [7] follows from this principle. If the flux function f is convex or concave between the left and right state of a discontinuity, then the solution forms either a shock or a rarefaction wave, i.e. a continuous function connecting the two states. Otherwise, combinations of shocks and rarefactions can result. These physical solutions are defined by a weak formulation of (1) accompanied by an entropy condition [7].

The first step in a particle method is to approximate the initial function u_0 by a finite number of points $x_1 \leqslant \cdots \leqslant x_m$ with function values $u_1, \ldots, u_m$. In Section 4, we present strategies on how to sample the initial function "well". The evolution of the solution is found by moving each particle x_i with speed $f'(u_i)$. This is possible as long as there are no "collisions" between particles. Two neighboring particles $x_i(t)$ and $x_{i+1}(t)$ collide at time $t + \Delta t_i$, where

$$\Delta t_i = -\frac{x_{i+1} - x_i}{f'(u_{i+1}) - f'(u_i)}. \tag{3}$$

A positive Δt_i indicates that the two particles at x_i and x_{i+1} will eventually collide. Thus, $t + \Delta t_s$ is the time of the next particle collision, where

$$\Delta t_s = \min\{\{\Delta t_i | \Delta t_i \geqslant 0\} \cup \infty\}. \tag{4}$$

For any time increment $\Delta t \leqslant \Delta t_s$ the particles can be moved directly to their new positions $x_i + f'(u_i)\Delta t$. Thus, we can step forward in time an amount Δt_s. Then, at least one particle will share its position with another. To proceed further, we merge each such pair of particles. If the collision time Δt_i is negative, the particles depart from each other. Although at each of the particles the correct function value is preserved, after some time their distance may be unsatisfyingly large, as the amount of error introduced during a merge grows with the size of the neighboring gaps. To avoid this, we insert new particles into large gaps (see Section 3.1) *before* merging particles.

In this paper, we present a method of merging and inserting particles in such a way that shocks move at correct speeds, and rarefactions have the correct shape. The strategy is based on mimicking the evolution of area for a conservation law, as is derived in Section 2. The definition of an area function gives rise to a natural interpolation between neighboring Lagrangian particles. As presented in Section 3, particle management can then be done to conserve area exactly. The resulting particle

method is shown to be TVD. Since the characteristic equation is solved exactly, and particle management is purely local, the method yields no numerical dissipation (where solutions are smooth) and correct shock speeds (where they are not). Specific strategies for sampling the initial data are discussed in Section 4.

In the remaining sections, the method is analyzed and generalized. In Section 5, we prove that the numerical solutions satisfy the Kružkov entropy condition, thus showing that the method yields entropy solutions for convex entropy functions. In Section 6, we present how non-convex flux functions can be treated. Strategies to include sources are presented in Section 7. In Section 8, we apply the method to examples and compare it to traditional finite volume methods using CLAWPACK [3]. Conclusions are drawn in Section 9, as well as possible applications and extensions of the method outlined.

2. Evolution of area for scalar conservation laws

Consider a one dimensional scalar conservation law

$$u_t + f(x,u)_x = 0, \quad u(x,0) = u_0(x). \tag{5}$$

Its characteristic equations [7]

$$\begin{cases} \dot{x} = f_u(x,u) \\ \dot{u} = -f_x(x,u) \end{cases} \tag{6}$$

yield the movement and change of function value of a particle. Let $u(x,t)$ be a solution of (5). The change of area between two *fixed* points x_1 and x_2 is solely given by the flux function f as

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,dx = f(x_1,u(x_1,t)) - f(x_2,u(x_2,t)) = -[f]_{x_1}^{x_2}. \tag{7}$$

In contrast, the change of area between two Lagrangian particles $(x_1(t),u_1(t))$ and $(x_2(t),u_2(t))$, i.e. points that *move* according to (6), is given by

$$\frac{d}{dt}\int_{x_1(t)}^{x_2(t)} u(x,t)\,dx = (f_u(x_2,u_2)u_2 - f(x_2,u_2)) - (f_u(x_1,u_1)u_1 - f(x_1,u_1)) = F(x_2,u_2) - F(x_1,u_1) = [F]_{(x_1,u_1)}^{(x_2,u_2)}, \tag{8}$$

where $F = f_u u - f$ is the Legendre transform of f. That is, f is a Hamiltonian of the dynamics (6), and F is a Lagrangian. Eq. (7) (respectively (8)) yields the change of area between two Eulerian (Lagrangian) points, only by knowing the flux f (the Lagrangian F) at the two points. Hence, in the same fashion as (7) can be used to construct a conservative fixed grid method, we use (8) to construct a conservative *particle* method.

Consider an *area value* $A_i(t)$ associated with each particle, such that $[A]_{x_i}^{x_{i+1}} = A_{i+1} - A_i$ is the area between x_i and x_{i+1}. Assume the values A_i are known at $t = 0$. Then we can find the areas at any time by solving the system arising from Eqs. (6 and 8)

$$\begin{cases} \dot{x}_i = f_u(x_i,u_i) \\ \dot{u}_i = -f_x(x_i,u_i) \\ \dot{A}_i = F(x_i,u_i). \end{cases} \tag{9}$$

Remark 1. While $\dot{f} = 0$ (since f is a Hamiltonian of the dynamics), in general $\dot{F} \neq 0$. However, if the flux function satisfies

$$f_{xu}f_u u - f_{uu}f_x u - f_x f_u = 0, \tag{10}$$

then $\dot{F} = 0$ (by the chain rule). Property (10) is satisfied for instance if $f = f(u)$ or $f(x,u) = \varphi(x)u^k$. If $\dot{F} = 0$, the evolution of area is particularly simple, namely A_i changes at a *constant* rate F_i.

2.1. Space-independent flux

Henceforth we only consider flux functions that are independent of the spatial variable, $f = f(u)$. Thus, by Remark 1, the area between two Lagrangian points changes linearly, as does the distance between them

$$\frac{d}{dt}\int_{x_1(t)}^{x_2(t)} u(x,t)\,dx = [F(u)]_{u_1}^{u_2}, \tag{11}$$

$$\frac{d}{dt}(x_2(t) - x_1(t)) = \dot{x}_2(t) - \dot{x}_1(t) = f'(u_2) - f'(u_1) = [f'(u)]_{u_1}^{u_2}. \tag{12}$$

If the two points x_1 and x_2 move at different speeds, then there is a time t_0 (which may be larger or smaller than t) at which they have the same position. This assumes that they remain characteristic points between t and t_0, i.e. they do not interact with shocks. At time t_0, the distance and the area between the two points vanish. From (11) and (12) we have that

$$\int_{x_1(t)}^{x_2(t)} u(x,t)\,\mathrm{d}x = (t-t_0)\cdot[F(u)]_{u_1}^{u_2} x_2(t) - x_1(t) = (t-t_0)\cdot[f'(u)]_{u_1}^{u_2}.$$

In short, the area between two Lagrangian points can be written as

$$\int_{x_1(t)}^{x_2(t)} u(x,t)\,\mathrm{d}x = (x_2(t)-x_1(t))\;a_f(u_1,u_2), \tag{13}$$

where $a_f(u_1,u_2)$ is the nonlinear average function

$$a_f(u_1,u_2) = \frac{[f'(u)u - f(u)]_{u_1}^{u_2}}{[f'(u)]_{u_1}^{u_2}} = \frac{\int_{u_1}^{u_2} f''(u)u\,\mathrm{d}u}{\int_{u_1}^{u_2} f''(u)\,\mathrm{d}u}. \tag{14}$$

If there is only one flux function, we drop the subscript, and simply write $a(u_1,u_2)$. The integral form shows that a is indeed an average of u, weighted by f''. The evolution of area (13) is independent of the specific solution, since by assumption we have excluded all solutions for which a shock would interact with either characteristic point. The following lemma describes some properties of the nonlinear average $a(\cdot,\cdot)$.

Lemma 2. *Let f be strictly convex in $[u_L,u_U]$, that is, $f''>0$ in (u_L,u_U). Then for all $u_1,u_2\in[u_L,u_U]$, the average (14) is...*

(1) *the same for f and $-f$;*
(2) *symmetric, $a(u_1,u_2)=a(u_2,u_1)$;*
(3) *an average, i.e. $a(u_1,u_2)\in(u_1,u_2)$, for $u_1\neq u_2$;*
(4) *strictly increasing in both u_1 and u_2; and*
(5) *continuous at $u_1=u_2$, with $a(u,u)=u$.*

Due to the first two properties, we can assume WLOG that $f''>0$ and $u_1\leqslant u_2$ whenever convenient.

Proof. We prove the claims in turn.

(1,2) Multiplying both numerator and denominator by -1 yields the proof:

$$a_f(u_1,u_2) = \frac{-\int_{u_1}^{u_2} f''(u)u\,\mathrm{d}u}{-\int_{u_1}^{u_2} f''(u)\,\mathrm{d}u} = \frac{\int_{u_1}^{u_2} -f''(u)u\,\mathrm{d}u}{\int_{u_1}^{u_2} -f''(u)\,\mathrm{d}u} = a_{-f}(u_1,u_2) = \frac{\int_{u_2}^{u_1} f''(u)u\,\mathrm{d}u}{\int_{u_2}^{u_1} f''(u)\,\mathrm{d}u} = a_f(u_2,u_1).$$

(3) We bound a from above:

$$a(u_1,u_2) = \frac{\int_{u_1}^{u_2} f''(u)u\,\mathrm{d}u}{\int_{u_1}^{u_2} f''(u)\,\mathrm{d}u} < \frac{u_2\int_{u_2}^{u_1} f''(u)\,\mathrm{d}u}{\int_{u_2}^{u_1} f''(u)\,\mathrm{d}u} = u_2.$$

A similar argument bounds a from below.

(4) We show that $a(u_1,u_2)$ is strictly increasing in the second argument. Let $u_1<u_2<u_3, u_i\in[u_L,u_U]$. Then

$$a(u_1,u_3) = \frac{\int_{u_1}^{u_2} f''(u)u\,\mathrm{d}u + \int_{u_2}^{u_3} f''(u)u\,\mathrm{d}u}{\int_{u_1}^{u_3} f''(u)\,\mathrm{d}u} = \frac{a(u_1,u_2)\int_{u_1}^{u_2} f''(u)\,\mathrm{d}u + a(u_2,u_3)\int_{u_2}^{u_3} f''(u)\,\mathrm{d}u}{\int_{u_1}^{u_3} f''(u)\,\mathrm{d}u}.$$

Due to property 3 we have that $a(u_1,u_2)<u_2<a(u_2,u_3)$. Thus

$$a(u_1,u_3) > \frac{a(u_1,u_2)\int_{u_1}^{u_2} f''(u)\,\mathrm{d}u + a(u_1,u_2)\int_{u_2}^{u_3} f''(u)\,\mathrm{d}u}{\int_{u_1}^{u_3} f''(u)\,\mathrm{d}u} = a(u_1,u_2).$$

A similar argument shows the result for the first argument.

(5) Since $u_1<a(u_1,u_2)<u_2$ for all $u_1\neq u_2$, we have (by the Sandwich Theorem) that

$$u = \lim_{u_1\to u} u_1 \leqslant \lim_{u_1,u_2\to u} a(u_1,u_2) \leqslant \lim_{u_2\to u} u_2 = u.$$

Therefore, $\lim_{u_1,u_2\to u} a(u_1,u_2) = u$. □

3. Interpolation and particle management

The time evolution of Eq. (1) is described by the characteristic movement of the particles (6). Particle management is an "instantaneous" operation (i.e. happening at constant time) that allows the method to continue stepping forward in time. It is designed to conserve area: The function value of an inserted or merged particle is chosen such that area is unchanged by

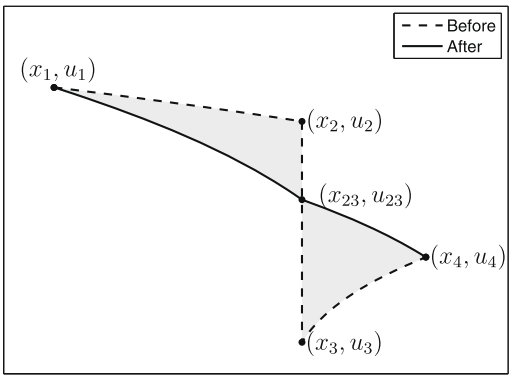

Fig. 1. Merging two particles yields a new particle with function value chosen "conservatively". This implies that the area of the two "triangles" is the same.

the operation. A simple condition guarantees that the entropy does not increase. In addition, we define an interpolating function between two neighboring particles, so that the change of area under the interpolating curve satisfies relation (11). This interpolation is shown to be an analytical solution of the conservation law.

3.1. Conservative particle management

Consider four neighboring particles located at $x_1 < x_2 \leqslant x_3 < x_4$ [1] with associated function values u_1, u_2, u_3, u_4. Assume that the flux f is strictly convex or concave on the range of function values $[\min_i(u_i), \max_i(u_i)]$. If $u_2 \neq u_3$, the particles' velocities must differ $f'(u_2) \neq f'(u_3)$, which gives rise to two possible cases that require particle management:

- **Inserting:** The two particles deviate, i.e. $f'(u_2) < f'(u_3)$. If $x_3 - x_2 \geqslant d_{\max}$ for some predefined maximum distance $d_{\max}$, we insert a new particle (x_{23}, u_{23}) with $x_2 < x_{23} < x_3$, such that the area is preserved:
$$(x_{23} - x_2)a(u_2, u_{23}) + (x_3 - x_{23})a(u_{23}, u_3) = (x_3 - x_2)a(u_2, u_3). \tag{15}$$
One can, for example, set $x_{23} = \frac{x_2+x_3}{2}$ and find u_{23} by (15), or set $u_{23} = \frac{u_2+u_3}{2}$ and find x_{23} by (15).
- **Merging:** The two particles collide, i.e. $f'(u_2) > f'(u_3)$. If $x_3 - x_2 \leqslant d_{\min}$ for some predefined minimum distance ($d_{\min} = 0$ is possible), we replace both with a new particle (x_{23}, u_{23}) with $x_2 \leqslant x_{23} \leqslant x_3$, such that the area is preserved:
$$(x_{23} - x_1)a(u_1, u_{23}) + (x_4 - x_{23})a(u_{23}, u_4) = (x_2 - x_1)a(u_1, u_2) + (x_3 - x_2)a(u_2, u_3) + (x_4 - x_3)a(u_3, u_4). \tag{16}$$

We choose $x_{23} = \frac{x_2+x_3}{2}$, and then find u_{23} such that (16) is satisfied. Fig. 1 illustrates the merging step.

Observe that inserting and merging are similar in nature. Conditions (15) and (16) for u_{23} are nonlinear (unless f is quadratic, see Remark 20). For most cases $u_{23} = \frac{u_2+u_3}{2}$ is a good initial guess, and the correct value can be obtained (up to the desired precision) by a few Newton iteration steps (or bisection, if the Newton iteration fails to converge). The next few claims attest that there is a unique value u_{23} that satisfies (15) and (16), respectively.

Lemma 3. *The function value u_{23} for the particle at x_{23} for Eqs. (15) and (16) is unique.*

Proof. From Lemma 2 we have that both $a(u_1, \cdot)$ and $a(\cdot, u_4)$ are strictly increasing. Thus, the LHS of both (15) and (16) are strictly increasing (in u_{23}), and cannot attain the same value for different values of u_{23}. □

Lemma 4. *There exists $u_{23} \in [u_2, u_3]$ which satisfies (15).*

Proof. WLOG we assume $u_2 \leqslant u_3$. We define
$$A = (x_3 - x_2)a(u_2, u_3),$$
$$B(u) = (x_{23} - x_2)a(u_2, u) + (x_3 - x_{23})a(u, u_3).$$
So Eq. (15) can be recast as $B(u_{23}) = A$. The monotonicity of $a(\cdot, \cdot)$ implies that $B(u_2) < A < B(u_3)$. Since a is continuous, so is B, and the result follows from the Intermediate Value Theorem. The proof for $u_2 > u_3$ is trivially similar. □

Lemma 5. *If $x_2 = x_3 = x_{23}$, there exists $u_{23} \in [u_2, u_3]$ which satisfies (16).*

[1] If more than two particles are at one position (x), all but the one with the smallest value (u) and the one with the largest value (u) are removed immediately.

Proof. The proof is identical to the proof of Lemma 4 with the following definition of A and $B(u)$:

$$A = (x_2 - x_1)a(u_1, u_2) + (x_4 - x_2)a(u_3, u_4),$$
$$B(u) = (x_2 - x_1)a(u_1, u) + (x_4 - x_2)a(u, u_4). \quad \square$$

Corollary 6. *If particles are merged and inserted according to Eqs. (15) and (16), then the total variation of the solution is either the same as before the operation, or smaller.*

Merging points only when $x_2 = x_3$ can be overly restrictive. The following theorem grants a little more freedom.

Theorem 7. *Consider four consecutive particles (x_i, u_i) $i = 1, \dots, 4$. If*

$$\frac{|u_3 - u_2|}{x_3 - x_2} \geqslant 4\left(\frac{\max|f''|}{\min|f''|}\right)^6 \frac{\max_i u_i - \min_i u_i}{\min(x_4 - x_3, x_2 - x_1)}, \tag{17}$$

then merging particles 2 and 3 with $x_{23} = \frac{x_2 + x_3}{2}$ yields $u_{23} \in [u_2, u_3]$.

The min and max of f'' are taken over the maximum range of $u_1, \dots, u_4$. Condition (17) is trivially satisfied if $x_2 = x_3$.

The idea of the proof is to consider merging in two steps. First, we find a value $\tilde{u}$ such that setting $u_2 = u_3 = \tilde{u}$ (while leaving x_2 and x_3 unchanged) preserves the area. Next, we merge the two particles to one with value u_{23} located at x_{23}. To prove the theorem we use two lemmas: Lemma 8 bounds $\tilde{u}$ *away* from u_2 and u_3 (but inside $[u_2, u_3]$). Lemma 9 bounds $|u_{23} - \tilde{u}|$ from above. We define three "area functions":

$$A = (x_2 - x_1)a(u_1, u_2) + (x_3 - x_2)a(u_2, u_3) + (x_4 - x_3)a(u_3, u_4),$$
$$B(u) = (x_2 - x_1)a(u_1, u) + (x_3 - x_2)a(u, u) + (x_4 - x_3)a(u, u_4),$$
$$C(u) = (x_2 - x_1)a(u_1, u) + (x_3 - x_2)\frac{1}{2}[a(u_1, u) + a(u, u_4)] + (x_4 - x_3)a(u, u_4).$$

Here A is the area before the merge that needs to be preserved, $B(u)$ is the area when the particles $2, 3$ have the value u, and $C(u)$ is the area when particles $2, 3$ have been merged to a single particle at $\frac{x_2 + x_3}{2}$ with value u.

Lemma 8. *The value $\tilde{u}$ for which $B(\tilde{u}) = A$ satisfies $\tilde{u} \in [u_2, u_3]$ and*

$$|\tilde{u} - u_i| \geqslant \frac{1}{2}\frac{\min(x_3 - x_1, x_4 - x_2)|u_3 - u_2|}{x_4 - x_1}\left|\frac{\min f''(u)}{\max f''(u)}\right|^4 \quad \text{for } i = 2, 3.$$

Lemma 9. *The value u_{23} for which $C(u_{23}) = A$ satisfies*

$$|u_{23} - \tilde{u}| \leqslant 2\frac{(x_3 - x_2)[\max(u_1, u_2, u_3, u_4) - \min(u_2, u_3)]}{x_4 - x_1}\left|\frac{\max f''(u)}{\min f''(u)}\right|^2.$$

The proofs of the last two lemmas are tedious and uninspiring; they are relegated to the appendix for the interested reader's perusal.

Proof (*of Theorem 7*). Starting from the hypothesis of the theorem, we find

$$\begin{aligned}\frac{1}{2}\frac{\min(x_4 - x_3, x_2 - x_1)|u_3 - u_2|}{x_4 - x_1}\left(\frac{\min|f''|}{\max|f''|}\right)^4 &\geqslant 2\left(\frac{\max|f''|}{\min|f''|}\right)^2 \frac{(\max_i u_i - \min_i u_i)(x_3 - x_2)}{x_4 - x_1} \\ &\geqslant 2\left(\frac{\max|f''|}{\min|f''|}\right)^2 \frac{(\max_i u_i - \min(u_2, u_3))(x_3 - x_2)}{x_4 - x_1}.\end{aligned}$$

Using Lemmas 8 and 9, we obtain

$$|\tilde{u} - u_i| \geqslant |u_{23} - \tilde{u}|$$

for $i = 2, 3$. Since we also have that $\tilde{u} \in [u_2, u_3]$ (from Lemma 9), we conclude that $u_{23} \in [u_2, u_3]$. $\square$

Remark 10. Due to Theorem 7, the merging step is robust with respect to small deviations in the distance of the merged particles. This also holds for the case $x_2 > x_3$, given the distance $|x_3 - x_2|$ is sufficiently small.

Theorem 11. *The particle method can run to arbitrary times.*

Proof. Let $u_L = \min_i u_i$, $u_U = \max_i u_i$, and $C = \max_{[u_L, u_U]}|f''(u)| \cdot (u_U - u_L)$. For any two particles, one has $|f'(u_{i+1}) - f'(u_i)| \leqslant C$. Define $\Delta x_i = x_{i+1} - x_i$. After each particle management, the next time increment (as defined in Section 1.1) is at least $\Delta t_s \geqslant \frac{\min_i \Delta x_i}{C}$. If we do not insert particles, then in each merge one particle is removed. Hence, a time evolution beyond any given time is possible, since the increments Δt_s will increase eventually. When a particle is inserted (whenever two

points are at a distance more than $d_{\max}$), the created distances are at least $\frac{d_{\max}}{2}$, preserving a lower bound on the following time increment. □

Remark 12 (*Choice of distance parameters*). If possible, the minimum particle distance should be chosen $d_{\min} = 0$. However, a small positive value of $d_{\min}$ also leads to a working method. This generalization is of interest if the characteristic equations have to be solved numerically, such as in Section 7.

The maximum particle distance $d_{\max}$ gives the minimal local resolution of the method. If the initial data is sampled with a resolution of h, then a good choice for the maximum distance is $h < d_{\max} < 2h$. Here, the upper bound comes from the fact that after an insertion the local particle distance is halved. Throughout our numerical simulations we use $d_{\max} = \frac{4}{3}h$ since this gives, on average, a distance of h between particles in a rarefaction. This is an estimate that results from solving $\frac{1}{2}(d_{\max} + \frac{1}{2}d_{\max}) = h$ for $d_{\max}$.

3.2. Conservative interpolation

Expression (13) defines an area between any two points. We show that this defines an interpolating function $v(x)$ between the two points. While an interpolation is not required for the particle management itself, it is useful for plotting the numerical solution, interpreting its properties, and including source terms. As derived in Section 2.1, the area between two points (x_1, u_1) and (x_2, u_2) equals

$$\int_{x_1}^{x_2} v(x)\,\mathrm{d}x = (x_2 - x_1)a(u_1, u_2),$$

assuming that f is strictly convex or concave in $[u_1, u_2]$. We define the interpolation by the principle that any point (x, v) on the function $v = v(x)$ must yield the same area when the interval is split:

$$(x - x_1)a(u_1, v) + (x_2 - x)a(v, u_2) = (x_2 - x_1)a(u_1, u_2). \tag{18}$$

If $u_1 = u_2$, the interpolant is a constant function. Otherwise, (18) can be rearranged to yield

$$\frac{x - x_1}{x_2 - x_1} = \frac{a(u_1, u_2) - a(u_2, v)}{a(u_1, v) - a(u_2, v)} = \frac{f'(v) - f'(u_1)}{f'(u_2) - f'(u_1)}. \tag{19}$$

The last equality in (19) follows from

Lemma 13. *For any u_1, u_2, u_3 with $u_1 \neq u_2$, the nonlinear average satisfies*

$$\frac{a(u_1, u_2) - a(u_2, u_3)}{a(u_1, u_3) - a(u_2, u_3)} = \frac{f'(u_3) - f'(u_1)}{f'(u_2) - f'(u_1)}.$$

Proof. By definition of the average function, we have

$$\begin{aligned}
0 &= a(u_1, u_2)[f'(u)]_{u_1}^{u_2} + a(u_2, u_3)[f'(u)]_{u_2}^{u_3} + a(u_3, u_1)[f'(u)]_{u_3}^{u_1} \\
&= a(u_1, u_2)[f'(u)]_{u_1}^{u_2} + a(u_2, u_3)\Big([f'(u)]_{u_1}^{u_3} - [f'(u)]_{u_1}^{u_2}\Big) - a(u_1, u_3)[f'(u)]_{u_1}^{u_3} \\
&= (a(u_1, u_2) - a(u_2, u_3))[f'(u)]_{u_1}^{u_2} - (a(u_1, u_3) - a(u_2, u_3))[f'(u)]_{u_1}^{u_3}.
\end{aligned}$$

Rearranging the terms proves the claim. □

Observe that condition (18) is identical to the condition for particle insertion (15), which means that any newly inserted particle must be placed on the interpolant. Relation (19) defines the interpolant as a function $x(v)$. This is in fact an advantage, since at a discontinuity $x_1 = x_2$, characteristics for all intermediate values v are defined. Thus, rarefaction fans arise naturally if $f'(u_1) < f'(u_2)$. If f has no inflection points between u_1 and u_2, the inverse function $v(x)$ exists. For plotting purposes we plot $x(v)$ instead of inverting the function.

The interpolation (19) can also be derived as a similarity solution of the conservation law (2), as follows. If $u_1 = u_2$, we define $v(x) = u_1$. Otherwise, one has $f'(u_1) \neq f'(u_2)$. As derived in Section 2, the solution either came from a discontinuity (i.e. it is a rarefaction wave) or it will become a shock (i.e. it is a compression wave). The time Δt_1 until this discontinuity happens is given by (3). At time $t + \Delta t_1$ the particles have the same position $x_1 = x_2 = x_{\mathrm{sh}}$, as shown in Fig. 2. At this time the interpolation must be a straight line connecting the two particles, representing a discontinuity at x_{sh}. We require any particle of the interpolating function $(x, v(x))$ to move with its characteristic velocity $f'(v(x))$ in the time between t and $t + \Delta t_1$. This defines the interpolation uniquely as

$$x(v) = x_1 - \Delta t_1(f'(v) - f'(u_1)) = x_1 + \frac{f'(v) - f'(u_1)}{f'(u_2) - f'(u_1)}(x_2 - x_1), \tag{20}$$

which equals expression (19).

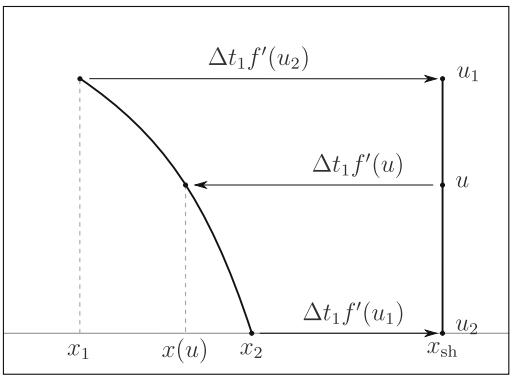

Fig. 2. To find the x-value of a particle with given u-value, one locates the shock and then travels to the current time with velocity $f'(u)$ as given by (20).

Not only is this interpolation compatible with the evolution of area under a conservation law, it also is a solution:

Lemma 14. *Together with the characteristic motion of the nodes, interpolation (20) is a solution of the conservation law (5).*

Proof. Using that $\dot{x}_i(t) = f'(u_i)$ for $i = 1, 2$ one obtains

$$\frac{\partial x}{\partial t}(v,t) = \dot{x}_1 + \frac{f'(v) - f'(u_1)}{f'(u_2) - f'(u_1)}(\dot{x}_2 - \dot{x}_1) = f'(u_1) + \frac{f'(v) - f'(u_1)}{f'(u_2) - f'(u_1)}(f'(u_2) - f'(u_1)) = f'(v).$$

Thus every point on the interpolation $v(x,t)$ satisfies the characteristic Eq. (6). □

Corollary 15 (Exact solution property). *Consider characteristic particles with $x_1(t) < x_2(t) < \ldots < x_n(t)$. At any time consider the function defined by the interpolation (20). This function is a classical (i.e. continuous) solution to the conservation law (5). In particular, it satisfies the conservation properties given in Section 2.*

This corollary breaks down when shocks occur.

Theorem 16 (TVD). *With the assumptions of Theorem 7, the particle method is total variation diminishing, thus it does not create spurious oscillations.*

Proof. Due to Corollary 15, the characteristic movement yields an exact classical solution, thus the total variation is constant. Particle insertion simply refines the interpolation, thus preserves the total variation. Due to Theorem 7, merging yields a new particle with a function value u_{23} between the values of the removed particles. Thus, the total variation is the same as before the merge or smaller. □

Remark 17. The particle method approximates the solution locally by similarity solutions, very similar to *front tracking* by Holden et al. [12]. Front tracking uses shocks (after approximating the flux function by a piecewise linear, and the solution by a piecewise constant function). In comparison, our method uses wave solutions, i.e. rarefactions and compressions.

3.3. Computational aspects

Remark 18 (*Shock location*). The particle method does not track shocks. Still, shocks can be located. Whenever particles are merged, the new particle can be marked as a *shock particle*. Thus, any shock stretches over three particles $(x_1,u_1),(x_2,u_2),(x_3,u_3)$, with the shock particle in the middle. Before plotting or interpreting the solution, a postprocessing step can be performed: The shock particle is replaced by a discontinuity, represented by two particles $(\bar{x}_2,u_1),(\bar{x}_2,u_3)$, with their position $\bar{x}_2$ chosen, such that area is conserved exactly. This step is harmless, since an immediate particle merge would recover the original configuration. As Fig. 3 illustrates, this postprocessing locates the shock with second order accuracy. The numerical results in Section 8.1 second this.

Remark 19 (*Shock speed*). Since the particle method locates shocks with second order accuracy (Lemma 18), the average shock speed is recovered with the same accuracy. However, the instantaneous speed of a reconstructed shock is only a first order accurate approximation to the true shock speed. The reason is that the shock position is chosen according to a locally constant solution. This yields an $O(h)$ error in the local shock height, and thus an $O(h)$ error in the shock speed, according to the Rankine–Hugoniot condition [7]. In addition, at each particle management, the position of the reconstructed shock jumps a distance of order $O(h^2)$. Note that Riemann problems are solved *exactly* by our method.

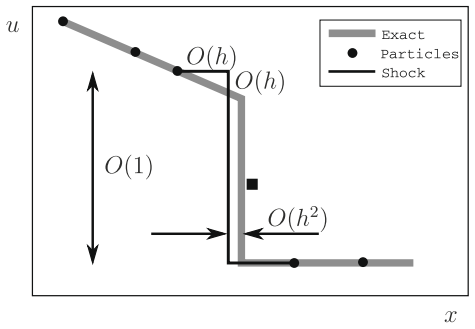

Fig. 3. The location of the reconstructed shock is second order accurate in h. The square particle is a "shock particle".

Remark 20 (*Quadratic flux function*). The method is particularly efficient for quadratic flux functions. In this case the interpolation (20) between two points is a straight line, and the average (14) is the arithmetic mean $a(u_1,u_2)=\frac{u_1+u_2}{2}$. Thus, the function values for particle management can be computed explicitly.

Remark 21 (*Computational cost*). An interesting aspect arises in the computational cost of the method, when counting evaluations of the flux function f and its derivatives f',f''. Particle movement does not involve any evaluations, since the characteristic velocity of each particle $f'(u_i)$ does not change. Consider a solution on $t\in[0,1]$ with a bounded number of shocks, to be approximated by $O(n)$ particles. Every particle merge and insertion requires $O(1)$ evaluations. After each time increment (4), $O(1)$ management steps are required. The total number of time increment steps is $O(n)$. Thus, the total cost is $O(n)$ evaluations, as opposed to $O(n^2)$ evaluations for Godunov methods. Note that this aspect is only apparent if evaluations of f',f'' are expensive, since the total number of operations is still $O(n^2)$.

4. Sampling of the initial data

When no source terms are present, the method has two sources of error: the initial sampling, and the merging of particles which contributes to the error in the neighborhood of shocks. Since, after the initial sampling and away from shocks, the solution is evolved exactly, it is natural to look for ways to reduce the error due to the initial sampling. In some applications, the initial function u_0 may be representable exactly by the interpolation (20). In other cases, it has to be approximated. In Section 4.1, we show how well u_0 can be approximated, as more and more particles are used. In Section 4.2, we outline a strategy of initializing a given number of particles in order to obtain a good approximation.

4.1. Error convergence

Lemma 22. *Consider a smooth function $w(x)$ on an interval $x\in\left[-\frac{h}{2},\frac{h}{2}\right]$, with $|f''(\bar w)|\geqslant C>0\ \forall\bar w\in\left[w\left(-\frac{h}{2}\right),w\left(\frac{h}{2}\right)\right]$. Let $v(x)$ denote the interpolant (20) between $-\frac{h}{2}$ and $\frac{h}{2}$. Then $|v(x)-w(x)|=O(h^2)$, and $\int_{-h/2}^{h/2}|v(x)-w(x)|\,dx=O(h^3)$ i.e. the approximation is second order accurate in both L^∞ and L^1 norms.*[2]

Proof. Substituting a Taylor expansion $w(x)=w_0+w_0'x+\frac{1}{2}w_0''x^2+O(x^3)$ into f' yields

$$f'(w(x))=f'(w_0)+f''(w_0)w_0'x+\frac{1}{2}\left(f''(w_0)w_0''+f'''(w_0){w_0'}^2\right)x^2+O(x^3). \tag{21}$$

Using (21) in (20) yields for the interpolation $v(x)$

$$\begin{aligned}f'(v(x))&=f'\left(w\left(-\frac{h}{2}\right)\right)+\left(f'\left(w\left(\frac{h}{2}\right)\right)-f'\left(w\left(-\frac{h}{2}\right)\right)\right)\frac{x+\frac{h}{2}}{h}\\&=f'(w_0)-\frac{1}{2}f''(w_0)w_0'h+\frac{1}{8}\left(f''(w_0)w_0''+f'''(w_0){w_0'}^2\right)h^2+O(h^3)+\left(f''(w_0)w_0'h+O(h^3)\right)\left(\frac{x}{h}+\frac{1}{2}\right)\\&=f'(w_0)+f''(w_0)w_0'x+\frac{1}{8}\left(f''(w_0)w_0''+f'''(w_0){w_0'}^2\right)h^2+O(h^3).\end{aligned} \tag{22}$$

Comparing (21) and (22) yields

[2] The power 3 in the order of the integral is needed so that the global L^1 error, which results from adding up the errors from all the intervals, is second order.

$$f'(w(x)) - f'(v(x)) = \frac{1}{2}\left(f''(w_0)w_0'' + f'''(w_0){w_0'}^2\right)\left(x^2 - \frac{1}{4}h^2\right) + O(h^3). \tag{23}$$

Using the Mean Value Theorem, we obtain the estimate

$$w(x) - v(x) = \frac{1}{2}\frac{\left(f''(w_0)w_0'' + f'''(w_0){w_0'}^2\right)\left(x^2 - \frac{1}{4}h^2\right)}{f''(\bar{w})} + O(h^3) \tag{24}$$

for some function value $\bar{w}$ between v and w (or $w(-\frac{h}{2})$ and $w(\frac{h}{2})$). Since the denominator is bounded from below by C, we have our result for the L^∞ norm. The L^1 error of the interpolation in the interval satisfies

$$\lim_{h\to 0}\frac{\int_{-h/2}^{h/2}|w(x) - v(x)|\,\mathrm{d}x}{h^3} = \frac{1}{12}\frac{\left(f''(w_0)w_0'' + f'''(w_0){w_0'}^2\right)}{f''(w_0)}. \tag{25}$$

Thus completing the proof. □

The formula for the error "density" is used for our adaptive sampling in the next section. We therefore name this density e:

$$e(x) = \frac{1}{12}\frac{\left(f''(w(x))w''(x) + f'''(w(x))w'^2(x)\right)}{f''(w(x))}. \tag{26}$$

Non-smooth functions can be approximated as well, if the discontinuities are known:

Theorem 23. *Consider a piecewise smooth function $u_0(x)$ with finitely many discontinuities (at known locations). Assume further that $|f''(u_0(x))| \geqslant C > 0\ \forall x$. Then u_0 can be approximated with second order accuracy, using the interpolations (20).*

Proof. First, represent each discontinuity in u_0 exactly, using two particles. This consumes only a finite number of points, thus the asymptotic behavior is not influenced. Second, place the remaining particles equidistantly at $(x_i, u_0(x_i))$. Since the jumps are represented exactly, the maximum error is second order by Lemma 22. □

For non-convex flux functions, presented in Section 6, the flux function can have inflection points at particles. In an interval bounded by an inflection particle, the second order accuracy is, in general, lost. First order accuracy is guaranteed by the following

Lemma 24. *Consider a smooth function $w(x)$ on an interval $x \in [0, h]$, with $|f''(\bar{w})| > 0 \forall \bar{w} \in (w(0), w(h))$ (so that the interpolation (20) exists). Then the interpolation $v(x)$ between 0 and h, given by (20), is at least first order accurate.*

Proof. The interpolation $v(x)$ is monotonous, hence one can bound

$$|v(x) - w(x)| \leqslant |w_{\min} - w_{\max}| \leqslant Ch,$$

where $C \geqslant \max_{x\in[0,h]}|w'(x)|$ and $w_{\min}$, and $w_{\max}$ are the minimum and maximum values that $w(x)$ attains over the interval. □

Remark 25. If the initial function is such that the flux crosses an inflection point, the error attains the form $\|v - w\|_{L^1} \sim ah^2 - b\log(h)h^2$ as $h \to 0$. The local L^1 error in a single interval abutting the inflection point is of order $O(h^2)$. The remaining intervals add a total L^1 error of order $O(h^2\log(h))$. Hence, the approximation is less than second order, but greater than first order.

4.2. Adaptive sampling

Due to Theorem 23, the initial data can be approximated with second order accuracy using equidistantly spaced points. Yet, for a fixed number of points, a non-equidistant spacing can yield a better approximation. The presented particle method is designed for non-equidistant points. Hence, adaptive sampling strategies can be easily used.

To minimize the error for a given number of points we use a particle density proportional to $e^{-\frac{1}{2}}(x)$. For further information on this see [5], for example. We define the integral

$$E(x) = \int_0^x e^{\frac{1}{2}}(\xi)\,\mathrm{d}\xi.$$

and sample $n+1$ points at positions

$$x_i = E^{-1}\left(E(L)\frac{i}{n}\right).$$

In the example presented in Fig. 6, this type of adaptive sampling is shown to reduce the initial error by a factor of about 2.

5. Entropy

We have argued in Section 3.2 that due to the constructed interpolation the particle method naturally distinguishes shocks from rarefaction fans. In this section, we show that the method in fact satisfies the entropy condition

$$\eta(u)_t + q(u) \leqslant 0 \tag{27}$$

for a convex entropy function η, if the shocks are resolved well enough during a merge step. The following lemma considers the Kružkov entropy pair $\eta_k(u) = |u-k|$, $q_k(u) = \operatorname{sign}(u-k)(f(u)-f(k))$. Holden and Risebro [13, Chapter 2.1] show that if (27) is satisfied for η_k, q_k (for all k), then it is satisfied for any convex entropy function. Using the interpolation (20) we show that the numerical solution obtained by the particle method satisfies this condition.

Lemma 25 (Entropy for merging). *Let $x_1 < x_2 = x_3 < x_4$ be the locations of four particles, with particles 2 and 3 to be merged, and $f'' > 0$[3], i.e. $u_2 > u_3$. If the value u_{23} resulting from the merge satisfies $u_1 \geqslant u_{23} \geqslant u_4$, then the Kružkov entropy $\int |u-k|dx$ does not increase due to the merge.*

Proof. Consider the segment $[x_1, x_4]$. Let $u(x)$ and $\hat{u}(x)$ denote the interpolating function before and after the merge, respectively. The interpolating function u is monotone in the value of its endpoints, thus $u(x) \leqslant \hat{u}(x)$ for $x \in [x_2, x_4]$, and $u(x) \geqslant \hat{u}(x)$ for $x \in [x_1, x_2]$. The function

$$I_+(x) = \begin{cases} x & x > 0 \\ 0 & x \leqslant 0 \end{cases}$$

can be used to write $|x| = x + 2I_+(-x)$. We identify two possible cases: $k \leqslant u_{23}$ and $k \geqslant u_{23}$. In the first case, $k \leqslant u_{23}$, we write

$$\int_{x_1}^{x_4} |u-k|\,\mathrm{d}x = \int_{x_1}^{x_4} (u-k)\,\mathrm{d}x + 2\int_{x_1}^{x_4} I_+(k-u)\,\mathrm{d}x.$$

Due to the definition of $\hat{u}$ we have

$$= \int_{x_1}^{x_4} (\hat{u}-k)\,\mathrm{d}x + 2\int_{x_1}^{x_2} I_+(k-u)\,\mathrm{d}x + 2\int_{x_2}^{x_4} I_+(k-u)\,\mathrm{d}x.$$

Since $k \leqslant u$ on $[x_1, x_2]$ and that $I_+(u)$ is non-decreasing, we get

$$\geqslant \int_{x_1}^{x_4} (\hat{u}-k)\,\mathrm{d}x + 0 + 2\int_{x_2}^{x_4} I_+(k-\hat{u})\,\mathrm{d}x.$$

Using $k \leqslant \hat{u}$ on $[x_1, x_2]$ we replace the zero with a different integral

$$= \int_{x_1}^{x_4} (\hat{u}-k)\,\mathrm{d}x + 2\int_{x_1}^{x_2} I_+(k-\hat{u})\,\mathrm{d}x + 2\int_{x_2}^{x_4} I_+(k-\hat{u})\,\mathrm{d}x = \int_{x_1}^{x_4} (\hat{u}-k)\,\mathrm{d}x + 2\int_{x_1}^{x_4} I_+(k-\hat{u})\,\mathrm{d}x = \int_{x_1}^{x_4} |\hat{u}-k|\,\mathrm{d}x.$$

The other option we have is $k \geqslant u_{23}$. In this case the proof is quite similar, but we start with $|k-u|$ instead, and remember that on $[x_1, x_2]$ we have that $u \geqslant \hat{u}$:

$$\begin{aligned}
\int_{x_1}^{x_4} |k-u|\,\mathrm{d}x &= \int_{x_1}^{x_4} (k-u)\,\mathrm{d}x + 2\int_{x_1}^{x_4} I_+(u-k)\,\mathrm{d}x = \int_{x_1}^{x_4} (k-\hat{u})\,\mathrm{d}x + 2\int_{x_1}^{x_2} I_+(u-k)\,\mathrm{d}x + 2\int_{x_2}^{x_4} I_+(u-k)\,\mathrm{d}x \\
&\geqslant \int_{x_1}^{x_4} (k-\hat{u})\,\mathrm{d}x + 2\int_{x_1}^{x_2} I_+(\hat{u}-k)\,\mathrm{d}x + 0 = \int_{x_1}^{x_4} (k-\hat{u})\,\mathrm{d}x + 2\int_{x_1}^{x_2} I_+(\hat{u}-k)\,\mathrm{d}x + 2\int_{x_2}^{x_4} I_+(\hat{u}-k)\,\mathrm{d}x \\
&= \int_{x_1}^{x_4} (k-\hat{u})\,\mathrm{d}x + 2\int_{x_1}^{x_4} I_+(\,\hat{u}-k)\,\mathrm{d}x = \int_{x_1}^{x_4} |k-\hat{u}|\,\mathrm{d}x.
\end{aligned}$$

This ends the proof. □

The assumption of Lemma 25 implies that shocks must be reasonably well resolved before the points defining it are merged. It is satisfied if the points to the left and right of a shock points are not too far. The condition can be ensured by an *entropy fix*: a merge is rejected*a posteriori* if the resolution condition is not satisfied. Then, points are inserted near the shock, and the merge is re-attempted.

Remark 26. With the entropy fix, a merge does not necessarily reduce the number of points. Based on numerical evidence, we conjecture that the statement of Theorem 11 remains valid, although its proof cannot be transferred in a straightforward fashion.

Theorem 24. *The presented particle method yields entropy solutions.*

[3] For the case $f'' < 0$, all following inequality signs must be reversed.

Proof. During the characteristic movement of the points, the entropy is constant. This is due to Corollary 15 which tells that the interpolation is a classical solution to the conservation law. Particle insertion does not change the interpolation, thus it does not change the entropy. Merging does not increase the entropy if the conditions of Lemma 25 are satisfied. □

6. Non-convex flux functions

So far we have only considered flux functions without inflection points (i.e. f'' always has the same sign) on the range of function values. In this section, we generalize our method for flux functions f where f'' has a finite number of zero crossings $u'_1 < \cdots < u'_k$. Between two successive points $u \in [u'_i, u'_{i+1}]$ the flux function is either convex or concave. We impose the following requirement for any set of particles: between any two particles for which f'' has opposite sign, there must be an *inflection particle* (x, u'_i). Thus, between two neighboring particles, f never has an inflection point, and the fundamental ideas from the previous sections transfer. In particular, the characteristic movement of particles is unaffected, and the interpolation between two particles remains uniquely defined by (20). It has infinite slope at inflection points (since $f''(u'_i) = 0$), but this is mostly harmless. However, two complications arise. First, every proof that relies on having a lower bound on f'' does not transfer easily. Second, merging particles when an inflection particle is involved requires a special treatment. The standard approach, as presented in Section 3.1, removes two colliding points and replaces them with a point of a different function value. If an inflection particle is involved in a collision, points must be merged in a different way so that an inflection particle remains.

We present one such special merge for dealing with a single inflection point (we do not consider here the interaction of two inflection points). Also, for simplicity, we consider a collision with identical point positions. Since the inflection particle must remain (although its position may change), we consider five neighboring particles and not four as before. Let $(x_i, u_i), i = 1, \ldots, 5$ be these particles so that $x_2 = x_3$, $f''(u_3) = 0$, and (WLOG) $f''' > 0$, i.e. the inflection particle is the slowest. The other cases are simple symmetries of this situation. The special merge consists of three different attempts to find the new particle configuration. The first two attempts may fail to provide a solution, in which case the next is attempted.

(1) Remove particle 2 and increase x_3 such that area is preserved. Accept, if x_3 is not increased beyond x_4.
(2) Remove particle 2, set $x_3 = x_4$ and increase both such that area is preserved. Accept, if x_3 and x_4 are not increased beyond x_5.
(3) Remove particle 4, set $x_3 = x_5$ and lower u_2 such that area is preserved.

Fig. 4 provides a visual description of these three configurations.

Theorem 29. *One of the three attempts listed above will succeed.*

Proof. Following from continuity and monotonicity of the average function $a(\cdot,\cdot)$, the three steps provide a continuous, monotonous increase in area. In the first attempt, the smallest area is achieved with x_3 unchanged. This area is necessarily smaller than the original area (since one can also get here by lowering u_2 to u_3). The area increases as x_3 is increased. The configuration with $x_3 = x_4$ has the maximum area for the first attempt, and the minimum area for the second. Again, the area increases as $x_3 = x_4$ increase. The configuration with $x_3 = x_4 = x_5$ has the maximum area for the second attempt, and the minimum for the third. Area increases as the new value of u_2 increases, and achieves its maximum value for an unchanged u_2. This area is necessarily larger than the original area. Consequently, one of the attempts must succeed. □

Remark 30. The resulting configuration may involve a new discontinuity (since $x_3 = x_4$ or $x_3 = x_5$). However, this is not a shock, but a rarefaction, since the particles will move away from each other. Consequently, these particles should *not* be merged.

The five-point particle management guarantees that in each merging step one particle is removed, as used in Theorem 11. In Section 8.2, numerical results on the Buckley–Leverett equation are presented. Since each of the three attempts covers a non-overlapping range of areas, the resulting configuration is independent of the order in which they are attempted.

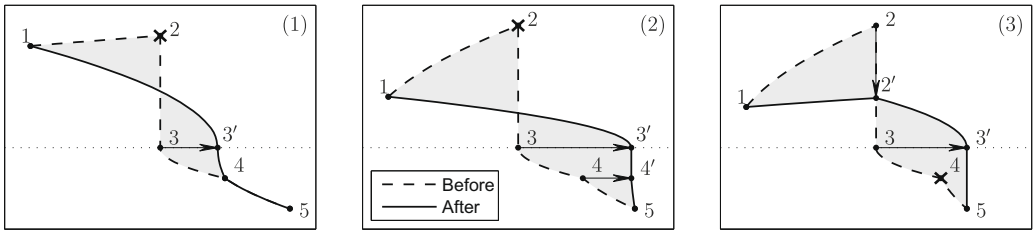

Fig. 4. Particle management around an inflection particle ($f''(u_3) = 0$) results in one of three possible configurations. Each configuration allows for more area under the function than the previous one. Here we see three archetypal particle configurations that result.

7. Sources

An important extension of the conservation law (1) is to allow a source term in the right hand side. This can be a function of x, t, the function value u, and in principle also of derivatives u_x, u_{xx}, etc. In the current work we consider the simple balance law

$$u_t + f(u)_x = g(x,u). \tag{28}$$

The method of characteristics [7] yields an evolution for each particle

$$\begin{cases} \dot{x} = f'(u), \\ \dot{u} = g(x,u). \end{cases} \tag{29}$$

With sources, the equation ceases to have exact conservation properties. Consequently, the interpolation derived in Section 3.2 is no longer a solution. While in special cases more complicated interpolation functions could be defined (depending on both f and g), here we construct an approximate method that is more general. Assume that the advection dominates over the source, which is the case in many applications. Thus, the interpolation and particle management are based solely on the flux function f.

The source g results in a vertical movement of the particles during their Lagrangian evolution. While in the absence of sources the next time of a particle merge can be computed a priori, now we solve the particle evolution (29) numerically, for instance by an explicit Runge–Kutta scheme. Merging takes place when two particles are too close (see Remark 10). In Section 8.3, we present numerical results.

Remark 31. The balance law (28) is solved correctly at characteristic points. Particle management, however, is based on an "incorrect" interpolation, since the source is neglected for the definition of area. The numerical results in Section 8.3 indicate that this does not cause problems for merging particles. However, inserting particles into large gaps may lead to significant misplacements, when the source is "active". Thus, with sources, insertion should either be avoided completely, or particles be adaptively refined. We shall address the important aspect of adaptivity in future work.

The presented approach incorporates sources directly into the characteristic equations. An alternative approach is operator splitting: First move particles neglecting the source, then correct function values according to the source. While the characteristic method is more precise, the splitting approach is more general. In particular, it can deal with source terms that involve derivatives of u.

8. Numerical results

The presented particle method is applied to various examples. In all cases, the "exact" reference solution is obtained or verified by a high resolution CLAWPACK [3] computation. We compare the accuracy of the particle method with numerical solutions obtained by CLAWPACK, considering similar resolutions. By construction, the particle method does not keep a fixed resolution. To compare resolutions we use the same number of particles to initialize the particle method as the number of cells in the corresponding CLAWPACK run. By keeping $d_{\max} = \frac{4}{3}h$, we find that the number of particles remains more-or-less constant throughout the computation. Shocks are located via post-processing before the error is measured.

In Section 8.1, the evolution and the formation of shocks of smooth initial data under a convex flux function are considered. The convergence error before and after the occurrence of shocks is investigated numerically. In Section 8.2, as an example of a non-convex flux function, the Buckley–Leverett equation is considered, and in Section 8.3, Burgers' equation with a source is simulated. The source code and all presented examples can be found on the `particleclaw` web page [19].

8.1. Convergence error

Fig. 5 shows the smooth initial function $u_0(x) = 0.5 + 0.2\exp(-x^2)\cos(\pi x)$, and its time evolution under the flux function $f(u) = \frac{1}{4}u^4$. Initially, we sample points on the function u_0. At time $t = 1$, the solution is still smooth, thus the particles lie exactly on the solution. By the time $t = 10$, a shock has emerged and interacted with a rarefaction. Although the numerical solution uses only a few points, it represents the true solution well.

From this example, the numerical accuracy of the particle method is extracted. For a sequence of particle densities, the initial data are sampled twice: equally spaced and adaptively. The particle method is applied with post-processing, as described in Remark 18. The error is measured in the true L^1 norm for function, which is possible due to the interpolation (20). Fig. 6 shows results for initial sampling error, and error after a time evolution. Initially ($t = 0$), the approximation is second order accurate for both sampling strategies (see Theorem 23). The advantage of the adaptive sampling is evident from the lower error that it creates in the interpolation.

After shocks have occurred ($t = 10$), the approximate solution without locating shocks is only first order accurate, since at any shock an error of the order height $\times$ width of the shock is made. However, the post-processing step recovers the second order accuracy. Hence, the particle method is second order accurate, even at locating shocks. One also sees that the advantage gained initially from the adaptive sampling is nearly lost at $t = 10$.

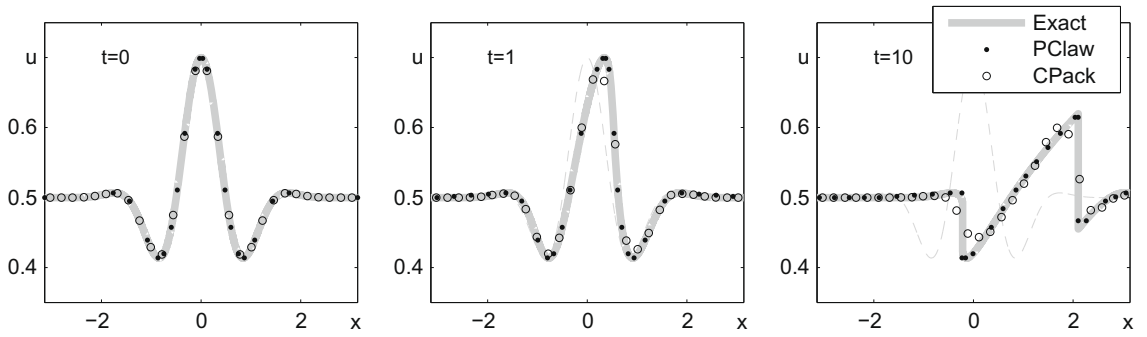

Fig. 5. The particle method for $f(u) = \frac{1}{4}u^4$ before and after shocks arise. The gray solid line is the hi-resolution solution from CLAWPACK and the dashed line is the initial condition.

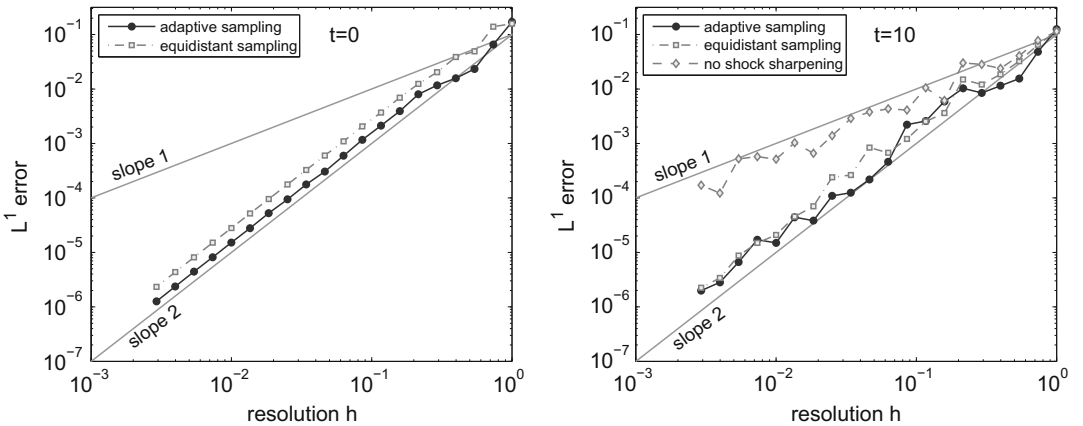

Fig. 6. L^1 Error of the particle method when solving for the flux function $\frac{1}{4}u^4$, initially, and after a shock has been formed. The figures compare equidistant with adaptive sampling and show the error that would result from not sharpening the shock.

For this example, CLAWPACK yields results of similar accuracy, as shown in Fig. 7. Since CLAWPACK is a method for calculating cell averages, we cannot find the true L^1 error. Instead, given a coarse-grid calculation and a fine-grid reference solution, we calculate the error in the area (function value times cell-size) for each of the coarse cells using the fine-grid solution. Adding all these errors together gives the relevant L^1 error. One can see that the CLAWPACK solution drops to first order accuracy around shocks, which is due to numerical dissipation. To investigate the accuracy away from shocks, we also consider the error while ignoring a small fixed domain surrounding each shock. The same error measure is also applied to the error calculations of our particle method. Of course, since post-processing already yields second order accuracy, this only reduces the size of the error, and does not change the order of convergence, as it does with CLAWPACK.

8.2. Non-convex flux function

As an example of a non-convex flux function, we consider the Buckley–Leverett equation

$$u_t + (f(u))_x = 0, \quad \text{with } f(u) = u^2 \Big/ \left(u^2 + \frac{1}{2}(1-u)^2\right), \tag{30}$$

which is a simple model for two-phase fluid flow in a porous medium (see LeVeque [15]). We consider piecewise constant initial data with a large downward jump crossing the inflection point, and a small upward jump. The large jump develops a shock at the bottom and a rarefaction at the top, the small jump is a pure rarefaction. Around $t = 0.2$, the two similarity solutions interact, thus lowering the separation point between shock and rarefaction. Fig. 8 shows numerical results. The solution obtained by the particle method (dots) is compared to a second order CLAWPACK solution (circles) of similar resolution. The particle method captures the behavior of the solution better; in particular, the rarefaction is represented very accurately. Only directly near the shock are inaccuracies visible. The solution away from the shock is nearly unaffected by the error at the shock.

Numerical results show (see Fig. 9) that both CLAWPACK and the presented particle method do not achieve second order accuracy for this problem. Nevertheless, the particle method has a much better accuracy than CLAWPACK. The drop in accu-

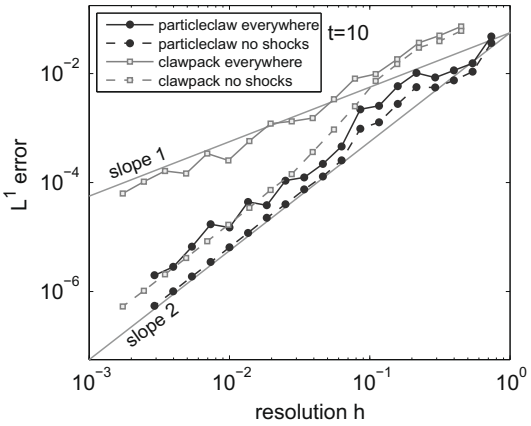

Fig. 7. A comparison of the errors given by the particle method and CLAWPACK when solving for the flux function $\frac{1}{4}u^4$. Without removing the errors from the shock region, CLAWPACK is only first order accurate.

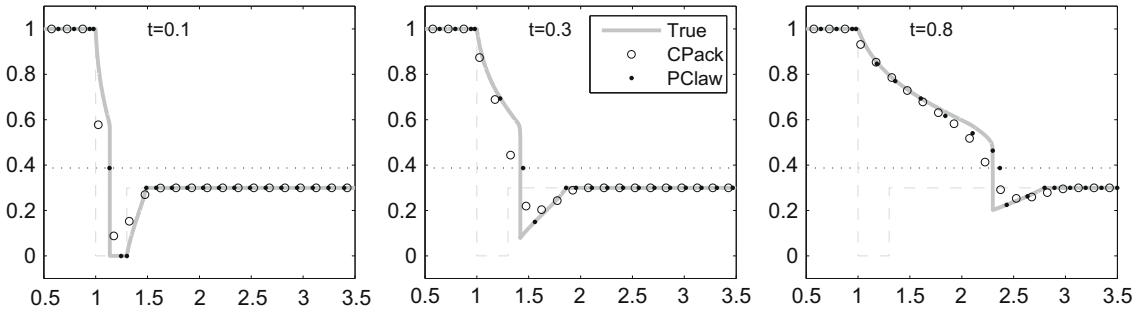

Fig. 8. Numerical results for the Buckley–Leverett equation at various times t.

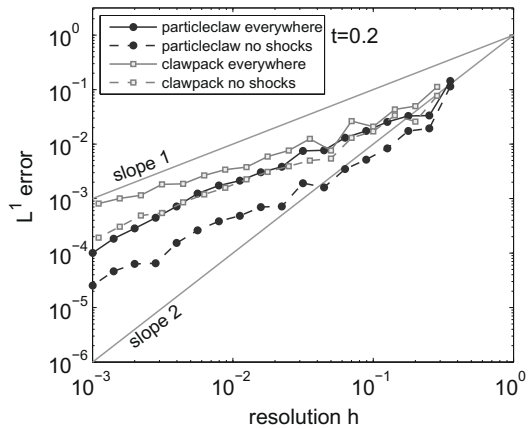

Fig. 9. A comparison of the errors given by the particle method and CLAWPACK when solving the Buckley–Leverett Eq. (30). While the particle method gets significantly better results than CLAWPACK (with or without the errors from around the shock), both methods are less than second order accurate.

racy is, presumably, due to inflection point in the Buckley–Leverett flux function, similar to the drop in accuracy of the sampling outlined in Remark 25.

8.3. Source terms

We consider Burgers' equation with a source

$$u_t + \left(\frac{1}{2}u^2\right)_x = b'(x)u. \tag{31}$$

It is a simple model for shallow water flow over a bottom profile $b(x)$. As in [14], we consider the domain $x \in [0, 10]$, and choose

$$b(x) = \begin{cases} \cos(\pi x) & x \in [4.5, 5.5], \\ 0 & \text{otherwise}. \end{cases}$$

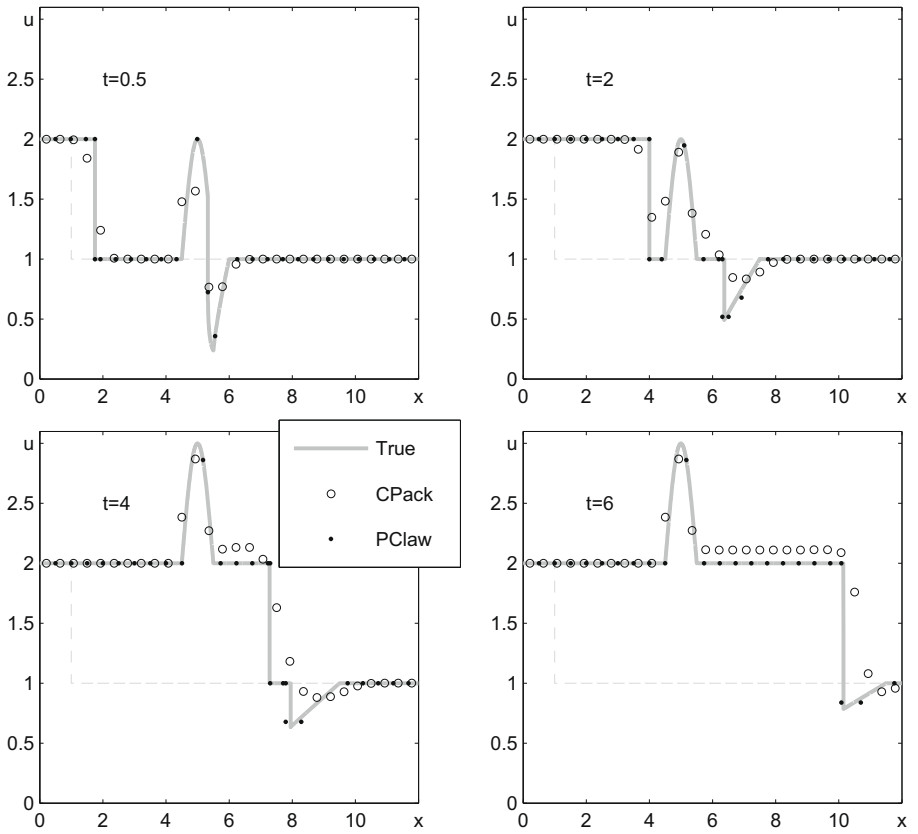

Fig. 10. Time evolution of Burgers' equation with a source as given by (31). The dots, circles, gray line and dashed line are, respectively, our particle method, CLAWPACK, the high resolution solution and the initial conditions.

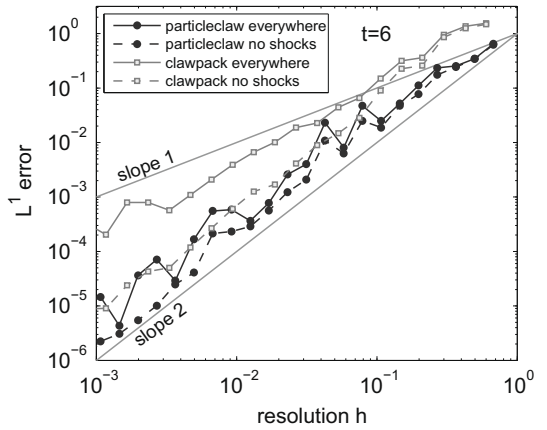

Fig. 11. A comparison of the errors given by the particle method and CLAWPACK when solving for Burgers' equation with a source term (31), at $t = 6$ (after the two shocks have interacted). Here too, the particle method provides a better accuracy than CLAWPACK.

The source term is included into the method of characteristics, as explained in Section 7. The time stepping is done by a fourth order Runge–Kutta scheme. Fig. 10 shows the computational results. The particle method (dots) approximates the solution significantly better than the second order CLAWPACK scheme (circles). Fig. 11 shows the error convergence of the particle method in comparison with CLAWPACK. One can observe that the particle method yields a smaller error than CLAWPACK. A particular aspect in favor of the characteristic approach is the precise (up to the resolution of the ODE solver) recovery of the function values after the obstacle. Since particles are moved independently according to the characteristic equations, an accurate time integration obtains the function values after the obstacle almost exactly, independent of the resolution of particles. Note that an efficient treatment of the source requires a special consideration of its discontinuities, either in the quadrature of the source (finite volume), or in the integration of the characteristic ODE (particle scheme).

9. Conclusions and outlook

We have presented a particle method that combines the method of characteristic, local similarity solutions, and particle management to a numerical scheme for 1D scalar conservation laws. The method conserves area exactly. It is TVD, yet second order accurate, even at locating shocks. It performs promisingly in various examples, as the numerical comparisons with a second order finite volume scheme show.

The particle method is an interesting alternative to fixed grid approaches, whenever conservation of mass is crucial, or shocks need to be located accurately. In addition, entropy is reduced only when particles are merged, which makes the approach suited for applications in which the evolution of mass and energy has to be reflected as precisely as possible. Furthermore, the method yields good results when few particles are used, in particular shocks between nearly-constant states are located well. This makes the approach attractive whenever scalar 1D conservation laws arise as sub-problems in a large computation, and only a few degrees of freedom can be devoted to the numerical solution of a single sub-problem. Examples are flows in networks (e.g. car traffic), and PDE constrained optimization.

As a first generalization, we have included source terms in the scheme. The method, still based on the method of characteristics, yields solutions of rather striking accuracy, compared to classical finite volume schemes. In future work, more general source terms will be considered, such as non-local convolutions, and terms involving derivatives of the solution. In these cases, the method of characteristics has to be replaced by a more general splitting approach.

Fundamental steps towards a more powerful particle method will be the generalization to higher space dimensions and to systems of conservation laws. Problems in multiple dimensions can be approximated by 1D problems using fractional steps. In this sense, the particle scheme could replace classical 1D Riemann solvers by 1D *wave solvers*. However, this approach is not fully satisfactory, since due to the required remeshing steps the benefits of a meshfree particle approach may be lost. On the other hand, with truly meshfree approaches in 2D/3D, one has to address the challenge that particles forming a shock need not necessarily collide. Possible remedies are the introduction of a numerical pressure, or the tracking of an unstructured triangular mesh. The movement of particles according to the method of characteristics can also be interpreted as a moving mesh approach [1]. Thus, ideas from this area could lead to particle strategies in higher space dimensions.

With systems, one difficulty is the presence of multiple characteristic velocities. One approach is to choose one Lagrangian velocity, which need not be a characteristic velocity. Coupling terms that appear in the moving frame equations are treated as source terms for each individual equation. Alternative approaches may use exact similarity solutions of the full system as building blocks. In this case, a single set of particles may not suffice, since two neighboring similarity solutions may interact.

Acknowledgments

The authors would like to thank R. LeVeque for helpful comments and suggestions. The support by the National Science Foundation is acknowledged. Y. Farjoun was partially supported by Grant DMS-0703937. B. Seibold was partially supported by Grant DMS-0813648.

Appendix A

The proofs of Lemmas 8 and 9 use a short lemma:

Lemma 32. *The derivative of $a(u,v)$ with respect to either of its variables is bounded from below and above as follows:*

$$\frac{1}{2}\left(\frac{\min f''}{\max f''}\right)^2 \leqslant \left[\frac{\partial a}{\partial u}(u,v), \frac{\partial a}{\partial v}(u,v)\right] \leqslant \frac{1}{2}\left(\frac{\max f''}{\min f''}\right)^2.$$

Here $\max f''$ *and* $\min f''$ *are taken over the interval* $[u,\ v]$.

Proof. This lemma follows from the definition of a:

$$\frac{\partial a}{\partial u}(u,v) = \frac{f''(u)\int_u^v f''(\omega)(\omega-u)\,\mathrm{d}\omega}{\left(\int_u^v f''(\omega)\,\mathrm{d}\omega\right)^2} \leqslant \left(\frac{\max f''}{\min f''}\right)^2 \frac{\int_u^v \omega - u\,\mathrm{d}\omega}{(v-u)^2} \leqslant \frac{1}{2}\left(\frac{\max f''}{\min f''}\right)^2.$$

The other bounds (on $\frac{\partial a}{\partial v}$ and the lower bound) have similar proofs. □

Proof (*of Lemma 8*). WLOG we assume that $u_3 \geqslant u_2$, and show for u_2. We bound $A - B(u_2)$ from below:

$$\begin{aligned} A - B(u_2) &= (x_3 - x_2)(a(u_2, u_3) - a(u_2, u_2)) + (x_4 - x_3)(a(u_3, u_4) - a(u_2, u_4)) \\ &\geqslant (x_3 - x_2)(u_3 - u_2) min \frac{\partial a}{\partial v}(u, v) + (x_4 - x_3)(u_3 - u_2) \min \frac{\partial a}{\partial u}(u, v) \geqslant (x_4 - x_2)(u_3 - u_2) \frac{1}{2} \left(\frac{\min f''}{\max f''} \right)^2. \end{aligned} \quad (A.1)$$

Since we are looking for $\tilde{u}$ such that $B(\tilde{u}) = A$, the previous bound is also a bound on $B(\tilde{u}) - B(u_2)$. From the Mean Value Theorem we have $\xi \in [\tilde{u},\ u_2]$ for which

$$\tilde{u} - u_2 = \frac{B(\tilde{u}) - B(u_2)}{B'(\xi)} = \frac{B(\tilde{u}) - B(u_2)}{(x_2 - x_1)\frac{\partial a}{\partial u}(\xi, u_1) + (x_4 - x_3)\frac{\partial a}{\partial u}(\xi, u_4) + (x_3 - x_2)} \geqslant \frac{B(\tilde{u}) - B(u_2)}{(x_4 - x_1)\left(\frac{\max f''}{\min f''}\right)^2}.$$

In the last step we used the upper bound on $\frac{\partial a}{\partial u}$ and that $\frac{\max f''}{\min f''} \geqslant 1$. From (A.1) we conclude that

$$\tilde{u} - u_2 \geqslant \frac{1}{2} \frac{(x_4 - x_2)(u_3 - u_2)}{x_4 - x_1} \left(\frac{\min f''}{\max f''} \right)^4.$$

Similarly, one can show that $u_3 - \tilde{u} \geqslant \frac{1}{2} \frac{(x_3 - x_1)(u_3 - u_2)}{x_4 - x_1} \left(\frac{\min f''}{\max f''} \right)^4$. □

Proof (*of Lemma 9*). Again, WLOG we assume that $u_3 \geqslant u_2$. This time we first bound $|C(\tilde{u}) - A|$ from above:

$$|C(\tilde{u}) - A| = C(\tilde{u}) - B(\tilde{u}) = \frac{x_3 - x_2}{2}(a(u_1, \tilde{u}) + a(\tilde{u}, u_4) - 2a(\tilde{u}, \tilde{u})) \leqslant (x_3 - x_2)[\max(u_i) - \min(u_2, u_3)].$$

Recall that $C(u_{23}) = A$, and that $C' > 0$ due to the monotonicity of a. Thus, for some point ξ

$$|\tilde{u} - u_{23}| = \frac{|C(\tilde{u}) - C(u_{23})|}{C'(\xi)} \leqslant \frac{(x_3 - x_2)[\max(u_i) - \min(u_2, u_3)]}{\min C'} \leqslant 2 \frac{(x_3 - x_2)[\max(u_i) - \min(u_2, u_3)]}{(x_4 - x_1)} \left(\frac{\max f''}{\min f''} \right)^2. \quad \square$$

References

[1] M.J. Baines, M.E. Hubbard, P.K. Jimack, A moving mesh finite element algorithm for the adaptive solution of time-dependent partial differential equations with moving boundaries, Appl. Numer. Math. 54 (2005) 450–469.
[2] A.J. Chorin, Numerical study of slightly viscous flow, J. Fluid Mech. 57 (1973) 785–796.
[3] Clawpack, Website, <http://www.clawpack.org>.
[4] R. Courant, E. Isaacson, M. Rees, On the solution of nonlinear hyperbolic differential equations by finite differences, Commun. Pure Appl. Math. 5 (1952) 243–255.
[5] L. Devroye, Non-uniform Random Variate Generation, Springer, New York, 1986.
[6] G.A. Dilts, Moving least squares particles hydrodynamics Im consistency and stability, Int. J. Numer. Methods Eng. 44 (1999) 1115–1155.
[7] L.C. Evans, Partial differential equations, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, 1998.
[8] Y. Farjoun, B. Seibold, Solving one dimensional scalar conservation laws by particle management, in: M. Griebel, M.A. Schweitzer (Eds.), Meshfree Methods for Partial Differential Equations IV, Lecture Notes in Computational Science and Engineering, vol. 65, Springer, 2009, pp. 95–109.
[9] R.A. Gingold, J.J. Monaghan, Smoothed particle hydrodynamics – theory and application to nonspherical stars, Mon. Not. Roy. Astron. Soc. 181 (1977) 375.
[10] S.K. Godunov, A difference scheme for the numerical computation of a discontinuous solution of the hydrodynamic equations, Math. Sbornik 47 (1959) 271–306.
[11] A. Harten, B. Engquist, S. Osher, S. Chakravarthy, Uniformly high order accurate essentially non-oscillatory schemes. III, J. Comput. Phys. 71 (2) (1987) 231–303.
[12] H. Holden, L. Holden, R. Hegh-Krohn, A numerical method for first order nonlinear scalar conservation laws in one dimension, Comput. Math. Appl. 15 (6–8) (1988) 595–602.
[13] H. Holden, N.H. Risebro, Front Tracking for Hyperbolic Conservation Laws, Springer, 2002.
[14] K.H. Karlsen, S. Mishra, N.H. Risebro, Well-balanced schemes for conservation laws with source terms based on a local discontinuous flux formulation, Math. Comput. 78 (2009) 55–78.
[15] R.J. Le Veque, Finite Volume Methods for Hyperbolic Problems, first ed., Cambridge University Press, 2002.
[16] X.-D. Liu, S. Osher, T. Chan, Weighted essentially non-oscillatory schemes, J. Comput. Phys. 115 (1994) 200–212.
[17] L. Lucy, A numerical approach to the testing of the fission hypothesis, Astron. J. 82 (1977) 1013–1024.
[18] J.J. Monaghan, Smoothed particle hydrodynamics, Rep. Prog. Phys. 68 (8) (2005) 1703–1759.
[19] Particleclaw, Website, <http://www-math.mit.edu/seibold/research/particleclaw/>.
[20] B. van Leer, Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme, J. Comput. Phys. 14 (1974) 361–370.